

-15-

Serial No.: 09/529,232

Respectfully submitted,

Date: December 10, 2003

By:

Christian Cawthorn, Reg. No. 47,352  
Agent of record

**OGILVY RENAULT**  
1981 McGill College, Suite 1600  
Montréal (Québec) Canada, H3A 2Y3

Tel: (514) 847-4256

Enc.

**CERTIFICATE OF FACSIMILE TRANSMISSION**

I hereby certify that this paper is being facsimile transmitted to the  
Patent and Trademark Office on the date shown below.

Christian Cawthorn

Name of person signing certification



Signature

December 10, 2003

Date

PTO/SB/01 (12-97)

**COMBINED  
DECLARATION FOR UTILITY OR DESIGN  
PATENT APPLICATION (37 CFR 1.63)  
AND POWER OF ATTORNEY**

Declaration Submitted with Initial Filing      OR       Declaration Submitted after Initial Filing (surcharge (37 CFR 1.16(e)) required)

Attorney Docket Number	
First Named Inventor	Yasuo Konishi et al.
Complete if known	
Application Number	
Filing Date	
Group Art Unit	
Examiner Name	

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe that I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

**TRIVALENT THROMBIN INHIBITOR**

the specification of which

is attached hereto.

OR

was filed on 10/15/1997  
(mm/dd/yyyy)

as United States Application Number or PCT International Application Number  
and was amended on \_\_\_\_\_ (if applicable).

PCT/CA97/00745

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment specifically referred to above.

I acknowledge the duty to disclose information which is material to patentability as defined in 37 CFR 1.56.

I hereby claim foreign priority benefits under 35 U.S.C. 119(a)-(d) or 365(b) of any foreign application(s) for patent or inventor's certificate, or 365(a) of any PCT international application which designated at least one country other than the United States of America, listed below and have also identified below, by checking the box, any foreign application for patent or inventor's certificate, or of any PCT international application having a filing date before that of the application on which priority is claimed.

Prior Foreign Application Number(s)	Country	Foreign Filing Date (MM/DD/YYYY)	Priority Not claimed	Certified Copy Attached? YES	Certified Copy Attached? NO
			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Additional foreign application numbers are listed on a supplemental priority data sheet PTO/SB/02B attached hereto:

I hereby claim the benefit under 35 U.S.C. 119(e) of any United States provisional application(s) listed below.

Application Number(s)	Filing Date (MM/DD/YYYY)	<input type="checkbox"/> Additional provisional application numbers are listed on a supplemental priority data sheet PTO/SB/02B attached hereto.
		<input type="checkbox"/>

**COMBINED DECLARATION FOR UTILITY OR DESIGN  
PATENT APPLICATION (37 CFR 1.63) AND POWER OF ATTORNEY**

I hereby claim the benefit under 35 U.S.C. 120 of any United States application(s), or 365(c) of any PCT international application designating the United States of America, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International application in the manner provided by the first paragraph of 35 U.S.C. 112, I acknowledge the duty to disclose information which is material to patentability as defined in 37 CFR 1.56 which became available between the filing date of the prior application and the national or PCT International filing date of this application.

U.S. Parent Application or PCT Parent Number	Parent Filing Date (MM/DD/YYYY)	Parent Patent Number (if applicable)

Additional U.S. or PCT International application numbers are listed on a supplemental priority data sheet PTO/SB/02B attached hereto.

As a named inventor, I hereby appoint the following registered practitioner(s) to prosecute this application and to transact all business in the Patent Trademark Office connected therewith:

Customer Number: **020988**



**020988**

PATENT AND TRADEMARK OFFICE

Direct all correspondence to:



**020988**

PATENT AND TRADEMARK OFFICE

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under 18 U.S.C. 1001 and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

**Name of sole or First Inventor:**

A petition has been filed for this unsigned inventor

Given Name (first and middle [if any])

Family Name or Surname

Yasuo

KONISHI

Inventor's Signature

*Yasuo Konishi*

Date November 26, 2003

Residence: City

Kirkland

State

Quebec

Country

CANADA

Citizenship

Canadian

Post Office Address

17130 Brunswick Boulevard

City

Kirkland

Province  
or State

Quebec

Postal  
Or Zip

H9J 1KB

Country

CANADA

Additional Inventors are being named on the

supplemental Additional Inventor(s) PTO/SB/02A attached hereto.

COMBINED DECLARATION FOR UTILITY OR DESIGN  
PATENT APPLICATION (37 CFR 1.63) AND POWER OF ATTORNEY

PTO/SB/02A (10-00)

DECLARATIONADDITIONAL INVENTOR(S)  
Supplemental Sheet  
Page 1 of 1

## Name of Additional Joint Inventor, if any:

 A petition has been filed for this unsigned inventor

Given Name (first and middle [if any])

Family Name or Surname

Jacek

SLON

Inventor's Signature

*Jacek Slon*

Date

Nov 17 2003

Residence:

City Toronto State Ontario Country CANADA Citizenship CanadianPost Office Address 302 Silverstone Drive unit 17City TorontoProvince  
or State

Ontario

Postal  
Or ZipCode  
M9V 3J4

Country

CA

## Name of Additional Joint Inventor, if any:

 A petition has been filed for this unsigned inventor

Given Name (first and middle [if any])

Family Name or Surname

Inventor's Signature

Date

Residence:

City \_\_\_\_\_ State \_\_\_\_\_ Country \_\_\_\_\_ Citizenship \_\_\_\_\_

Post Office Address \_\_\_\_\_

City \_\_\_\_\_

Province  
or StatePostal  
Or Zip

Country \_\_\_\_\_

## Name of Additional Joint Inventor, if any:

 A petition has been filed for this unsigned inventor

Given Name (first and middle [if any])

Family Name or Surname

Inventor's Signature

Date

Residence:

City \_\_\_\_\_ State \_\_\_\_\_ Country \_\_\_\_\_ Citizenship \_\_\_\_\_

Post Office Address \_\_\_\_\_

City \_\_\_\_\_

Province  
or StatePostal  
Or Zip

Country \_\_\_\_\_

 Additional inventors are being named on the \_\_\_\_\_ supplemental Additional Inventor(s) PTO/SB/02A attached hereto.

## Design of P1' and P3' Residues of Trivalent Thrombin Inhibitors and Their Crystal Structures<sup>†</sup>

Jacek J. Slon-Usakiewicz, J. Sivaraman, Yunge Li, Miroslaw Cygler, and Yasuo Konishi\*

Biotechnology Research Institute, National Research Council of Canada, 6100 Royalmount Avenue, Montreal, Quebec H4P 2R2, Canada

Received October 18, 1999; Revised Manuscript Received December 14, 1999

**ABSTRACT:** Synthetic bivalent thrombin inhibitors comprise an active site blocking segment, a fibrinogen recognition exosite blocking segment, and a linker connecting these segments. Possible nonpolar interactions of the P1' and P3' residues of the linker with thrombin S1' and S3' subsites, respectively, were identified using the "Methyl Scan" method [Slon-Usakiewicz et al. (1997) *Biochemistry* 36, 13494-13502]. A series of inhibitors (4-*tert*-butylbenzenesulfonyl)-Arg-(D-pipeolic acid)-Xaa-Gly-Yaa-Gly- $\beta$ Ala-Asp-Tyr-Glu-Pro-Ile-Pro-Glu-Glu-Ala-( $\beta$ -cyclohexylalanine)-(D-Glu)-OH, in which nonpolar P1' residue Xaa or P3' residue Yaa was incorporated, were designed and improved the affinity to thrombin. Substitution of the P3' residue with D-phenylglycine or D-Phe improved the  $K_i$  value to  $(9.5 \pm 0.6) \times 10^{-14}$  or  $1.3 \pm 0.5 \times 10^{-13}$  M, respectively, compared to that of a reference inhibitor with Gly residues at Xaa and Yaa residues ( $K_i = (2.4 \pm 0.5) \times 10^{-11}$  M). Similarly, substitution of the P1' residue with L-norleucine or L- $\beta$ -(2-thienyl)alanine lowered the  $K_i$  values to  $(8.2 \pm 0.6) \times 10^{-14}$  or  $(5.1 \pm 0.4) \times 10^{-14}$  M, respectively. The linker Gly-Gly-Gly- $\beta$ Ala of the inhibitors in the previous sentence was simplified with 12-aminododecanoic acid, resulting in further improvement of the  $K_i$  values to  $(3.8 \pm 0.6) \times 10^{-14}$  or  $(1.7 \pm 0.4) \times 10^{-14}$  M, respectively. These  $K_i$  values are equivalent to that of natural hirudin ( $2.2 \times 10^{-14}$  M), yet the size of the synthetic inhibitors (2 kD) is only one-third that of hirudin (7 kD). Two inhibitors, with L-norleucine or L- $\beta$ -(2-thienyl)alanine at the P1' residue and the improved linker of 12-aminododecanoic acid, were crystallized in complex with human  $\alpha$ -thrombin. The crystal structures of these complexes were solved and refined to 2.1 Å resolution. The Lys<sup>60F</sup> side chain of thrombin moved significantly and formed a large nonpolar S1' subsite to accommodate the bulky P1' residue.

Thrombin is a proteolytic regulatory enzyme of coagulation cascade. Upon vascular injury, thrombin is activated from prothrombin and converts soluble fibrinogen into insoluble fibrin clot. Thrombin also hydrolyzes the amide bonds of protein C and thrombin receptor on the platelet surface. The new N-terminal sequences generated by hydrolysis of various substrates have been shown in several cases to have specific biological functions. For example, after release of fibrinopeptide A, the newly formed N-terminal Gly-Pro-Arg-Val- sequence of fibrin binds to the C-terminal domain of other fibrin molecules to form fibrin clot (1).<sup>1</sup> The newly formed N-terminal Ser<sup>42</sup>-Phe-Leu-Leu-Arg... of the thrombin receptor after cleaving at the Arg<sup>41</sup>-Ser<sup>42</sup> bond binds intramolecularly and activates the receptors of human platelets and vascular endothelial cells (2). Thus, the sequence of P' residues has dual functions, as a substrate and as a functional site of the cleaved protein.

Thrombin cleaves proteins at a limited number of sites. The primary specificity of thrombin is determined by P1 residue Arg/Lys due to an ionic interaction with Asp<sup>189</sup> at the bottom of the S1 subsite.<sup>2,3</sup> However, thrombin cleaves

only four Arg-Gly bonds among 376 Arg/Lys-Xaa bonds in fibrinogen to release fibrinopeptides (3). The high specificity is thus arrived from other interactions. The S2 subsite of thrombin is nonpolar, and Pro is by far the preferred P2 residue (6). The S3 subsite of thrombin is called "the aryl binding site" and prefers aromatic or nonpolar amino acids, e.g., Phe<sup>8</sup> and Leu<sup>9</sup> of fibrinogen, to occupy the S3 subsite (7). Thrombin also has unique insertion loops B (Tyr<sup>60A</sup>, Ile<sup>61B</sup>) and C (Thr<sup>149A</sup>-Ala<sup>149B</sup>), contributing to the substrate specificity. Especially, it has been pointed out that the Lys<sup>60F</sup> side chain excludes bulky amino acids of the P1' residue (4). For example, a mutagenesis of Lys<sup>60F</sup> with Ala enhanced the affinity of the bulky Leu P1' residue (8). The P2' residue in natural substrates and inhibitors of thrombin is very diverse, e.g., Phe, Leu, Val, Ile, Pro, Gly, His, Glu, and Asn; however,  $k_{cat}/K_m$  of synthetic substrates showed a preference for bulky hydrophobic P2' residues such as Phe and Trp.

<sup>2</sup> The substrate/inhibitor and subsite nomenclature was suggested by Schechter and Berger (3) such that residues of substrates/inhibitors were numbered P1, P2, P3, etc., toward the amino terminus from the scissile peptide bond, and the complementary subsites of the enzyme were numbered S1, S2, S3, etc., respectively. The residues of some inhibitors such as MD-805, NAPAP, and TAPAP are exceptionally numbered as P3, P1, and P2 from the N-terminus, reflecting their occupations of thrombin S3, S1, and S2 subsites, respectively, in the complex.

<sup>3</sup> The numbering of human  $\alpha$ -thrombin residues is based on the chymotrypsin sequence (4).

\* NRCC Publication No. 42940.

\* To whom correspondence should be addressed. Phone: (514) 496-6339. Fax: (514) 496-5143. E-mail: Yasuo.Konishi@nrc.ca.

<sup>1</sup> Amino acids used in this paper are in the L-configuration unless specified as D-amino acid for D-isomer and D,L for racemic mixture.

## P1' and P3' Residues of Thrombin Inhibitors

whereas acidic amino acids of Glu and Asp were not favored (6). Similarly, the P3' residue in natural substrates and inhibitors of thrombin is very diverse, e.g., Arg, Asn, His, Leu, Gly, and Asp; however,  $k_{cat}/K_m$  of synthetic substrates showed a preference for basic and bulky hydrophobic amino acid as the P3' residue, whereas acidic amino acids of Glu and Asp were not favored (6). Indeed, protein C, having Asp at the P3' residue, is a poor substrate of thrombin, and is actively hydrolyzed only when the substrate specificity of thrombin is changed by the binding of thrombomodulin. The S' subsites of thrombin form a long and deep cleft toward the FRE.<sup>4</sup> The physiological function of this cleft is not well understood.

Bivalent thrombin inhibitors bind to the active site and the FRE of thrombin simultaneously, and a linker connects these two blocking segments. Three types of bivalent inhibitors have been developed. The first type of inhibitor is a (substrate-type active site inhibitor segment [e.g., (D-Phe)<sup>H1</sup>-Pro<sup>H2</sup>-Arg<sup>H3</sup>-Pro<sup>H4</sup>]-linker)-(FRE inhibitor segment [e.g., hirudin<sup>55-65</sup>]) (9-11).<sup>5</sup> Since the scissile bond at Arg<sup>H3</sup>-Pro<sup>H4</sup> is slowly hydrolyzed by thrombin, various pseudopeptide bonds or transition-state analogues have been developed. However, the synthesis of the pseudopeptide bonds or transition-state analogues may require tedious organic chemistry with low yield (12-15). Inhibitors of the second type mimic the binding mode of hirudin with a typical sequence of (P1-P2-P3 or P2-P1-P3)-(linker)-(FRE inhibitor segment) (16-18). The active site inhibitor segment binds in reverse to the substrates, and a long linker bridges between the P3 residue and the N-terminal Asp<sup>H5</sup> of the FRE inhibitor segment. These inhibitors advantageously avoid the scissile bond or are able to incorporate transition-state analogues. This arrangement eliminated the N-terminal free amino group of the P3 residue, which, in the substrate-type inhibitor segment, enhanced the affinity ~100-fold (19). Thus, the affinity of inhibitors of this type could be only <10-fold better than the corresponding active site inhibitors (17). Inhibitors of the third type utilize high affinity of nonsubstrate-type inhibitors such as MD-805 and NAPAP (20). The scissile bond is eliminated by connecting the P1' residue directly to the P2 residue. Tsuda et al. (20) successfully designed this class of inhibitors with  $K_i$  values on the order of  $10^{-11}$  M. The inhibitors were synthesized by using conventional SPPS with ~70% purity in the crude products. The inhibitors also made it possible to design the linker as P' residues for additional interactions with thrombin since short linkers lie along the deep cleft at the S' subsites.

<sup>4</sup> Abbreviations:  $\alpha$ -Abu,  $\alpha$ -aminobutyric acid;  $\beta$ -Ala,  $\beta$ -alanine;  $\lambda$ -Add, 12-aminododecanoic acid;  $\alpha$ -Aib,  $\alpha$ -aminoisobutyric acid; AMC, 7-amino-4-methylcoumarin; Bbs, 4-*tert*-butylbenzenesulfonyl; Cha,  $\beta$ -cyclohexylalanine; Chg, cyclohexylglycine; Fmoc, 9-fluorenylmethoxycarbonyl; FRE, fibrinogen recognition exo site; Hph, homophenylalanine; HPLC, high-performance liquid chromatography; MD-805, (2R,4R)-4-methyl-1-[N $\alpha$ -(3-methyl-1,2,3,4-tetrahydro-8-quinolinesulfonyl)-L-arginyl]-2-piperidinecarboxylic acid; NAPAP, N $\alpha$ -(2-naphthylsulfonylglycyl)-D,L-p-aminophenylalanylpiperidine; Nle, norleucine; Nva, norvaline; Pen,  $\beta$ , $\beta$ -dimethylcysteine; Phe, phenylglycine; Pip, pipelic acid; SPPS, solid-phase peptide synthesis; Tbg, *tert*-butylglycine; TFA, trifluoroacetic acid; Thi,  $\beta$ -(2-thienyl)alanine; Tos, p-toluenesulfonyl; Tris, 2-amino-2-(hydroxymethyl)-1,3-propanediol.

<sup>5</sup> The numbering of the inhibitor residues is based on the hirudin sequence, and "H" is added to the residue number of the inhibitors, e.g., Bbs<sup>H1</sup>-Arg<sup>H2</sup>-(D-Pip)<sup>H3</sup>-Gly<sup>H4</sup>-Gly<sup>H5</sup>-Gly<sup>H6</sup>-Gly<sup>H7</sup>- $\beta$ -Ala<sup>H8</sup>-Asp<sup>H9</sup>-Tyr<sup>H10</sup>-Glu<sup>H11</sup>-Pro<sup>H12</sup>-Ile<sup>H13</sup>-Pro<sup>H14</sup>-Glu<sup>H15</sup>-Glu<sup>H16</sup>-Ala<sup>H17</sup>-Cha<sup>H18</sup>-(D-Glu)<sup>H19</sup>-OH.

## Biochemistry, Vol. 39, No. 9, 2000 2385

Nonpolar interactions were searched for using a "Methyl Scan" in which a methyl side chain was introduced at each backbone atom of the linker and the  $K_i$  values of the inhibitors were evaluated (21). The methyl side chain at the second and eighth atoms of the 16-atom linker, corresponding to the P1' and P3' residues, respectively, improved the affinity of the inhibitor by 20-25-fold with L- and D-configurations, respectively.<sup>6</sup> This was rather contradictory to the strong preference for small amino acids as the P1' residue. The preference for nonpolar D-amino acids at the P3' residue was novel since no D-amino acids had been studied.

In this paper a series of inhibitors, Bbs-Arg-(D-Pip)-Xaa-Gly-Yaa-Gly- $\beta$ -Ala-Asp-Tyr-Glu-Pro-Ile-Pro-Glu-Ala-Cha-(D-Glu)-OH, were used, where Xaa and Yaa are Gly or various D- or L-nonpolar amino acids. The crystal structures of two inhibitors in complex with thrombin are presented.

## EXPERIMENTAL PROCEDURES

**Materials.** Human  $\alpha$ -thrombin (3000 NIH units/mg) for enzymatic assay, bovine fibrinogen (70% protein, 85% of protein clottable), Tos-Gly-Pro-Arg-AMC-HCl salt, poly(ethylene glycol) 8000, and Tris were purchased from Sigma (St. Louis, MO). Human  $\alpha$ -thrombin for crystallization was purchased from Haematologic Technologies Inc. (Essex Junction, VT) and was used without further purification. All Fmoc-amino acids and all other amino acid derivatives for peptide synthesis were purchased from Advanced ChemTech (Louisville, KY), Bachem Bioscience Inc. (King of Prussia, PA), and Calbiochem-Novabiochem (San Diego, CA). Fmoc-D-Glu(OtBu)-Wang resin (0.59 mmol/g) was purchased from Calbiochem-Novabiochem. The solvents for peptide synthesis were obtained from B&J Chemicals (Muskegon, MI) and Perkin-Elmer/Applied Biosystems (Foster City, CA). Citric acid, diethyl ether, and glacial acetic acid were purchased from Anachemia Chemical Inc. (Rouses Point, NY). TFA was purchased from Halocarbon Products Co. (River Edge, NJ). Phenol, thioanisole, potassium sulfate, sodium phosphate, and ethanedithiol were purchased from Aldrich (Milwaukee, WI).

**Peptide Synthesis.** The peptides were synthesized by using the solid-phase method on a 396 multiple peptide synthesizer (Advanced ChemTech) using a conventional Fmoc procedure. Peptides were cleaved from the resin using Reagent K (TFA (82.5%)/water(5%)/phenol (5%)/thioanisole (5%)/ethanedithiol (2.5%); 25 mL/g peptide-resin) for 2-4 h at room temperature. After precipitation with diethyl ether, peptides were filtered, dissolved in 50% acetic acid, and lyophilized. The peptides were then purified by a preparative HPLC (Vydac C<sub>4</sub> column, 4.6  $\times$  25 cm) using a linear gradient of 20-50% acetonitrile in 0.1% TFA (0.5%/min gradient; 33 mL/min flow rate). The final products were lyophilized with 98% or higher purity estimated by an analytical HPLC (Vydac C<sub>18</sub> column, 0.46  $\times$  25 cm; 10-60% acetonitrile in 0.1% TFA; 1.0%/min gradient; 1.0 mL/

<sup>6</sup> The linkers were named according to the number of atoms contributing to the length of the linkers; e.g., a linker 4-aminobutyric acid is a 5-atom linker and (12-aminododecanoic acid-4-aminobutyric acid) is an 18-atom linker. For convenience, in this paper, the P1' moiety of the linker is connected to the P2 residue and marks the start of the linker. This differs from the previous numbering scheme where the linker began after the P1' residue (22).

min flow rate). The elution profile was monitored by an absorbance at 210 and 254 nm for the analytical HPLC, and 220 nm for the preparative HPLC. The peptides were identified with a Beckman model 6300 amino acid analyzer (Beckman, Fullerton, CA) and an API III mass spectrometer (Sciex, Concord, Ontario, Canada). Amino acid analysis was used for peptide content determination. All peptides used in this paper have correct amino acid compositions and molecular masses.

**Amidolytic Assays.** The inhibition of amidolytic activity of human  $\alpha$ -thrombin was measured fluorometrically using Tos-Gly-Pro-Arg-AMC as a fluorogenic substrate in 50 mM Tris-HCl buffer (pH 7.80 at 25 °C) containing 0.1 M NaCl and 0.1% poly(ethylene glycol) 8000 at room temperature (23). The final concentrations of the inhibitors, the substrate, and human  $\alpha$ -thrombin were  $(0.5-1000)K_i$ ,  $(1-8) \times 10^{-6}$  M, and  $3.0 \times 10^{-11}$  M, respectively, if  $K_i > 10^{-10}$  M;  $(10-100)K_i$ ,  $(5-40) \times 10^{-6}$  M, and  $3.0 \times 10^{-11}$  M, respectively, if  $10^{-10}$  M  $> K_i > 10^{-11}$  M; and  $(3.3-56) \times 10^{-10}$ ,  $(5-40) \times 10^{-6}$ , and  $3.0 \times 10^{-11}$  M, respectively, if  $K_i < 10^{-11}$  M. The hydrolysis of the substrate by thrombin was monitored on a Hitachi F2000 fluorescence spectrophotometer and a Parkin-Elmer LS50B luminescence spectrophotometer ( $\lambda_{ex} = 383$  nm;  $\lambda_{em} = 455$  nm), and the fluorescent intensity was calibrated by using AMC. The substrate and an inhibitor were premixed at appropriate concentrations (the solution volume was adjusted to 2.99 mL) before 10  $\mu$ L of human  $\alpha$ -thrombin ( $9 \times 10^{-2}$  M) was added. The reaction reached a steady state within 3 min after the hydrolysis started. The steady-state velocity was then measured for a few minutes. The kinetic data of the steady-state velocity at various concentrations of the substrate and the inhibitors of the competitive inhibition were analyzed using the methods described in refs 22 and 24. Microsoft Excel was used to estimate the kinetic parameters ( $K_m$ ,  $V_{max}$ , and  $K_i$ ).

**Crystallization and Data Collection.** Two thrombin inhibitors, P628 [Bbs-Arg-(D-Pip)-Nle- $\lambda$ Adod-Asp-Tyr-Glu-Pro-Ile-Pro-Glu-Glu-Ala-Cha-(D-Glu)-OH] and P798 [Bbs-Arg-(D-Pip)-Thi- $\lambda$ Adod-Asp-Tyr-Glu-Pro-Ile-Pro-Glu-Glu-Ala-Cha-(D-Glu)-OH], were crystallized in complex with human  $\alpha$ -thrombin. The thrombin-inhibitor complexes were prepared by overnight incubation of thrombin (6 mg/mL) with inhibitor P628 or P798 at a molar ratio of 1:2 in 50 mM potassium sulfate, 50 mM citrate buffer (pH 5.5) at 4 °C. The crystals were grown by using the hanging drop vapor diffusion method with a reservoir solution of 25% (w/v) poly(ethylene glycol) 4000, 40 mM potassium sulfate, and 100 mM citric acid buffered to pH 5.5 with sodium phosphate at 18 °C. The pH of the reservoir solution was adjusted. The crystallization drops contained 2  $\mu$ L of thrombin-inhibitor complex solution and 4  $\mu$ L of the reservoir solution. Rodlike crystals were harvested after 3–4 days. Diffraction data were collected on a Raxis II C area detector with a Rigaku RU 300 rotating anode source. Data evaluation of thrombin-P628 complex performed with Denzo and Scalepack (25, 26) established a  $C_2$  space group having cell parameters  $a = 71.1$  Å,  $b = 71.9$  Å,  $c = 73.3$  Å, and  $\beta = 100.5^\circ$ . A total of 85 433 reflections observed were merged to 20 926 unique reflections with an  $R_{sym} = 0.06$  and completeness of 98.1% to 2.1 Å resolution. Thrombin-P798 complex was also crystallized in the  $C_2$  space group having cell parameters  $a$

$= 70.8$  Å,  $b = 72.0$  Å,  $c = 72.9$  Å, and  $\beta = 100.2^\circ$ . A total of 70 010 reflections observed were merged to 20 426 unique reflections with an  $R_{sym} = 0.05$  and completeness of 95.8% to 2.1 Å resolution. Both crystals contained one complex molecule per asymmetric unit.

**Structure Determination and Refinement.** A thrombin molecule from the thrombin-hirutinin-2 complex (27) (PDB code 1lhs) was used as a model in the structural determination of thrombin-P628 and -P798 complexes. Program O version 5.10 (28) and X-plor 3.8 (29) were used in model building and in the refinement, respectively. Appropriate entries were added to the dictionaries of both programs to accommodate the nonstandard groups of the inhibitors. Rigid-body refinement of the thrombin model followed by minimization reduced the  $R$  factor to 0.32 for both complexes.

The  $3F_0 - 2F_c$  and difference maps calculated for thrombin at this stage showed clearly the inhibitors. They were modeled into these maps, except for D-Glu<sup>H65</sup>. One hundred cycles of energy minimization were followed by molecular dynamics refinement with slow cooling from 1500 to 300 K, and the cycle was completed by another 100 cycles of energy minimization. The individual  $B$  factors were refined in 20 cycles of  $B$  factor refinement. Solvent molecules were added to the model. Several cycles of the refinement were performed, followed by model refitting to  $3F_0 - 2F_c$ ,  $F_0 - F_c$ , and the omit maps. The final structure includes A-chain Asp<sup>1A</sup>-Lys<sup>14A</sup> and B-chain Ile<sup>16</sup>-Glu<sup>146</sup> and Gly<sup>150</sup>-Phe<sup>245</sup> of thrombin and all residues of P628 except D-Glu<sup>H65</sup> or all residues of P798 except Cha<sup>H64</sup>-(D-Glu<sup>H65</sup>). The thrombin-P628 and -P798 complexes also include 142 and 160 water molecules, respectively. The final  $R$  factor for the reflections within the 8–2.1 Å resolution range was 0.192 ( $R$ -free = 24.0%) and 0.187 ( $R$ -free = 23.3%) with  $I > \sigma(I)$  for the thrombin-P628 and -P798 complexes, respectively. The rms deviation of the bond length was 0.009 and 0.008 Å for the thrombin-P628 and -P798 complexes, respectively. The rms deviation of the bond angle was 1.72° and 1.70° for the thrombin-P628 and -P798 complexes, respectively. The backbone dihedral angles ( $\phi$  and  $\psi$ ) of the inhibitors and the thrombin residues were in the allowed region of the Ramachandra plot except Ala<sup>H63</sup> of P628.

## RESULTS AND DISCUSSION

**P1' Residue of the Bivalent Inhibitors.** A series of bivalent thrombin inhibitors Bbs-Arg-(D-Pip)-Xaa-Gly-Gly- $\beta$ Ala-Asp-Tyr-Glu-Pro-Ile-Pro-Glu-Glu-Ala-Cha-(D-Glu)-OH were synthesized by incorporating various nonpolar amino acids at the P1' residue Xaa, where Gly was used as a reference amino acid ( $K_i = 2.4 \times 10^{-11}$  M) (Table 1). As reported in Methyl Scan (21), the P1' residue Ala improved the affinity 20-fold with a  $K_i$  value of  $(1.2 \pm 0.4) \times 10^{-12}$  M (P1004), whereas D-Ala improved the affinity 6-fold with a  $K_i$  value of  $(4.2 \pm 0.5) \times 10^{-12}$  M (P1005), suggesting preference for L-amino acid at the P1' residue. Simultaneous incorporation of D- and L-Ala, i.e.,  $\alpha$ Aib, did not improve the affinity any further, with a  $K_i$  value of  $(2.4 \pm 0.5) \times 10^{-12}$  M (P1373), possibly due to the conformational constraint of the  $\alpha$ Aib backbone (30). Introduction of a  $\gamma$ -methyl group, i.e., substitution of the P1' residue Ala with  $\alpha$ Abu, improved the affinity 2-fold with a  $K_i$  value of  $(6.3 \pm 0.5) \times 10^{-13}$  M (P1378). Introduction of a  $\delta$ -methyl group by substituting

## "P1' and P3' Residues of Thrombin Inhibitors

Table 1: Inhibition Constant of Bivalent Thrombin Inhibitors  
Bbs-Arg-(D-Pip)-Xaa-Gly-Gly-Gly- $\beta$ Ala-Asp-Tyr-Glu-Pro-Ile-Pro-Glu-Glu-Ala-Cha-(D-Glu)-OH

ID no.	Xaa	$K_i$ (pM)	ID no.	Xaa	$K_i$ (pM)
P836	Gly	24 $\pm$ 5			
P1004	Ala	1.2 $\pm$ 0.4	P1005	D-Ala	4.2 $\pm$ 0.5
P1373	D- $\alpha$ Abu	2.4 $\pm$ 0.5			
P1378	D- $\alpha$ Abu	0.63 $\pm$ 0.05	P1379	D-D- $\alpha$ Abu	4.3 $\pm$ 0.4
P1376	Nva	0.24 $\pm$ 0.05	P1375	D-Nva	5.1 $\pm$ 0.4
P1377	Nle	0.082 $\pm$ 0.006	P1374	D-Nle	5.3 $\pm$ 0.3
P1387	Met	0.11 $\pm$ 0.03	P1395	D-Met	4.8 $\pm$ 0.3
P1381	Val	0.84 $\pm$ 0.05	P1380	D-Val	3.7 $\pm$ 0.4
P1420	Tbg	1.1 $\pm$ 0.3	P1421	D-Tbg	5.8 $\pm$ 0.4
P1438	Pen	1.5 $\pm$ 0.4	P1439	D-Pen	6.8 $\pm$ 0.5
P1383	Ile	0.14 $\pm$ 0.04	P1384	D-Ile	4.3 $\pm$ 0.3
P1382	Leu	0.12 $\pm$ 0.04	P1385	D-Leu	2.6 $\pm$ 0.5
P1399	Chg	0.35 $\pm$ 0.5			
P1388	Cha	0.12 $\pm$ 0.04	P1396	D-Cha	7.2 $\pm$ 0.3
P1397	Phg	3.1 $\pm$ 0.4	P1398	D-Phg	7.8 $\pm$ 0.5
P1386	Phe	0.51 $\pm$ 0.05	P1394	D-Phe	3.4 $\pm$ 0.3
P1400	Hph	0.18 $\pm$ 0.05	P1441	D-Hph	2.8 $\pm$ 0.5
P1391	Trp	630 $\pm$ 30	P1443	D-Trp	820 $\pm$ 50
P1389	His	0.91 $\pm$ 0.04	P1442	D-His	2.1 $\pm$ 0.3
P1392	Thi	0.051 $\pm$ 0.004	P1393	D-Thi	2.8 $\pm$ 0.4

the P1' residue D- $\alpha$ Abu with Nva further improved the affinity 2.5-fold with a  $K_i$  value of  $(2.4 \pm 0.5) \times 10^{-13}$  M (P1376). Addition of an  $\epsilon$ -methyl group by substituting the P1' residue Nva with Nle further improved the affinity 3-fold with a  $K_i$  value of  $(8.2 \pm 0.6) \times 10^{-14}$  M (P1377). Thus, moderate but steady 2-3-fold improvement of the affinity was observed when linear alkyl side chains were introduced at the P1' residue, except 20-fold drastic improvement when the methyl side chain of Ala was introduced. On the other hand, the corresponding D-amino acid substitutions with D- $\alpha$ Abu, D-Nva, and D-Nle showed no improvement. These linear alkyl side chains of D-amino acids might be exposed to the solvent or have little interaction with thrombin. Incorporation of Met, which is homologous to Nle, at the P1' residue showed an affinity, with a  $K_i$  value of  $(1.1 \pm 0.3) \times 10^{-13}$  M (P1387), similar to that ( $K_i = (8.2 \pm 0.6) \times 10^{-14}$  M) of Nle (P1377).

Branched analogues of  $\alpha$ Abu showed little effect on the affinity of the inhibitors, i.e., Val, Tbg, and Pen, with  $K_i$  values of  $(8.4 \pm 0.5) \times 10^{-13}$  (P1381),  $(1.1 \pm 0.3) \times 10^{-12}$  (P1420), and  $(1.5 \pm 0.4) \times 10^{-12}$  (P1438) M, respectively, compared to  $(6.3 \pm 0.5) \times 10^{-13}$  M of  $\alpha$ Abu (P1378). On the other hand, branched analogues of Nva showed weak improvements with  $K_i$  values of  $(1.4 \pm 0.4) \times 10^{-13}$  M with Ile (branching at the C $\beta$  atom) (P1383) and of  $(1.2 \pm 0.4) \times 10^{-13}$  M with Leu (branching at the C $\gamma$  atom) (P1382) compared to  $(2.4 \pm 0.5) \times 10^{-13}$  M of Nva (P1376). Chg and Cha are the bulky cyclic nonpolar side chains with reduced conformational flexibility. The  $K_i$  values of these analogues were  $(3.5 \pm 0.5) \times 10^{-13}$  (P1399) and  $(1.2 \pm 0.4) \times 10^{-13}$  (P1388) M for Chg and Cha, respectively, showing no improvement compared to the linear alkyl side chains. Among the linear, branched, and cyclic alkyl side chains, the inhibitor P1377 with Nle had the highest affinity. Since the amide bonds of the linker might affect the interactions of the P1' residue with thrombin, three amide bonds in the P1377 linker Gly-Gly-Gly- $\beta$ Ala were eliminated by substituting with  $\lambda$ Adod. The resulting inhibitor P628 [Bbs-Arg-(D-Pip)-Nle- $\lambda$ Adod-Asp-Tyr-Glu-Pro-Ile-Pro-Glu-Glu-Ala-Cha-(D-Glu)-OH] improved the affinity 2-fold to a

Table 2: Active Site Directed Thrombin Inhibitors  
Bbs-Arg-(D-Pip)-Xaa-NH<sub>2</sub>

ID no.	Xaa	$K_i$ (nM)	ID no.	Xaa	$K_i$ (nM)
P972	None	$180 \pm 20$	P898	Hph	$25 \pm 3$
P887	Gly	$550 \pm 30$	P899	Thi	$5.2 \pm 0.4$
P891	Nle	$25 \pm 4$			

$K_i$  value of  $(3.8 \pm 0.6) \times 10^{-14}$  M. The crystal structure of the thrombin-P628 complex is described below.

The aromatic side chain was also introduced at the P1' residue. Incorporation of Phg improved the affinity 8-fold with a  $K_i$  value of  $(3.1 \pm 0.4) \times 10^{-12}$  M (P1397) compared to that with Gly (P836) ( $(2.4 \pm 0.5) \times 10^{-13}$  M). The substitutions with longer aromatic side chains (Phe and Hph) improved the affinity further, i.e.,  $K_i = (5.1 \pm 0.5) \times 10^{-13}$  M for Phe (P1386) and  $(1.8 \pm 0.5) \times 10^{-13}$  M for Hph (P1400), suggesting a deep pocket of the S1' subsite. However, Trp was too bulky to fit at the S1' subsite, resulting in a high  $K_i$  value of  $(6.3 \pm 0.3) \times 10^{-10}$  M (P1391). The replacement of Phe with His (P1389) did not improve the affinity. However, when a His homologue (Thi) was in this position, the affinity was increased 10-fold, with  $K_i = (5.1 \pm 0.4) \times 10^{-14}$  M (P1392). When the Gly-Gly-Gly- $\beta$ Ala linker in this compound was replaced by  $\lambda$ Adod as in P628, the resulting inhibitor P798 [Bbs-Arg-(D-Pip)-Thi- $\lambda$ Adod-Asp-Tyr-Glu-Pro-Ile-Pro-Glu-Ala-Cha-(D-Glu)-OH] gained 3-fold in affinity,  $K_i = (1.7 \pm 0.4) \times 10^{-14}$  M. The crystal structure of the thrombin-P798 complex is described below. All examined inhibitors with D-amino acids at the P1' position showed a higher  $K_i$  value than that of the inhibitor with a corresponding L-amino acid (Table 1). Thus, the P1' position clearly shows preference for L-amino acids.

**P1' Residue of the Active Site Inhibitors.** The bivalent thrombin inhibitors in this paper use short 16-atom linkers. Since the 16-atom linkers are stretched along the deep cleft between the active site and the FRE (31), the P1' residue is forced to be in the vicinity of the S1' subsite. Such restriction was not present in the active site inhibitors Bbs-Arg-(D-Pip)-Xaa-NH<sub>2</sub> where various nonpolar amino acids were incorporated at the P1' residue Xaa (Table 2). A reference peptide, Bbs-Arg-(D-Pip)-Gly-NH<sub>2</sub> (P887), showed a  $K_i$  value of  $(5.5 \pm 0.3) \times 10^{-7}$  M. Bbs-Arg-(D-Pip)-NH<sub>2</sub> (P972), which is missing the P1' residue, showed a 3 times lower  $K_i$  value of  $(1.80 \pm 0.20) \times 10^{-7}$  M. Thus, the cost of introducing the P1' residue Gly is 0.66 kcal/mol. This cost may be overcome by incorporating an appropriate side chain at the P1' residue. The P1' amino acids which improved the affinity in the bivalent inhibitors, Bbs-Arg-(D-Pip)-Xaa-Gly-Gly- $\beta$ Ala-Asp-Tyr-Glu-Pro-Ile-Pro-Glu-Glu-Ala-Cha-(D-Glu)-OH, also improved the affinity of the active site inhibitors Bbs-Arg-(D-Pip)-Xaa-NH<sub>2</sub>. Nle improved the affinity of the active site inhibitor by 22-fold with a  $K_i$  value of  $(2.5 \pm 0.4) \times 10^{-8}$  M (P891) compared to the 290-fold improvement in the bivalent inhibitor (P1377). Similarly, Hph improved the  $K_i$  value by 22-fold ( $K_i = (2.5 \pm 0.3) \times 10^{-8}$  M) in the active site inhibitor (P898) whereas there was a 133-fold improvement in the bivalent inhibitor (P1400); Thi improved the  $K_i$  value 106-fold ( $K_i = (5.2 \pm 0.4) \times 10^{-9}$  M) in the active site inhibitor (P899) compared to a 470-fold improvement in the bivalent inhibitor (P1392). The smaller improvement by Nle, Hph, and Thi in the active site inhibitors compared to those in the bivalent inhibitors may come from the larger

Table 3: Inhibition Constant of Bivalent Thrombin Inhibitors  
Bbs-Arg-(D-Pip)-Gly-Gly-Yaa-Gly- $\beta$ Ala-Asp-Tyr-Glu-Pro-Ile-Pro-Glu-Glu-Ala-Cha-(D-Glu)-OH

ID no.	Yaa	K <sub>i</sub> (pM)	ID no.	Yaa	K <sub>i</sub> (pM)
P836	Gly	24 ± 5			
P1054	Ala	8.7 ± 0.2	P1055	D-Ala	0.96 ± 0.03
P1068	$\alpha$ Aib	1.4 ± 0.3			
P1056	$\alpha$ Abu	7.4 ± 0.3	P1346	D- $\alpha$ Abu	0.77 ± 0.03
P1057	Nva	9.2 ± 0.4	P1075	D-Nva	0.88 ± 0.04
P1058	Nle	8.9 ± 0.5	P1074	D-Nle	0.68 ± 0.04
P1418	Met	10.4 ± 0.5	P1419	D-Met	1.5 ± 0.5
P1061	Val	1.2 ± 0.5	P1070	D-Val	0.62 ± 0.02
P1069	Tbg	4.3 ± 0.5	P1073	D-Tbg	0.44 ± 0.04
P1431	Pen	5.6 ± 0.5	P1432	D-Pen	1.5 ± 0.5
P1059	Ile	1.9 ± 0.3	P1071	D-Ile	0.35 ± 0.5
P1060	Leu	2.2 ± 0.5	P1072	D-Leu	0.62 ± 0.03
P1341	Chg	3.2 ± 0.4			
P1424	Cha	9.6 ± 0.5	P1077	D-Cha	1.5 ± 0.5
P1340	Phg	2.8 ± 0.4	P1076	D-Phg	0.095 ± 0.006
P1339	Phc	4.5 ± 0.4	P1078	D-Phc	0.13 ± 0.03
P1345	Hph	7.2 ± 0.3	P1425	D-Hph	0.78 ± 0.05
P1344	Trp	9.8 ± 0.4	P1427	D-Trp	2.2 ± 0.5
P1343	His	6.2 ± 0.5	P1426	D-His	1.4 ± 0.4
P1342	Thi	5.1 ± 0.5	P1428	D-Thi	1.2 ± 0.4

entropy loss of the P1' residue in the active site inhibitors than that in the bivalent inhibitors.

**P3' Residue of the Bivalent Inhibitors.** The Methyl Scan of the linker showed a preference for D-amino acid at the P3' residue; i.e., substitutions of the P3' residue Gly by L- and D-Ala improved the affinity 3- and 25-fold, respectively (27). A series of bivalent thrombin inhibitors Bbs-Arg-(D-Pip)-Gly-Gly-Yaa-Gly- $\beta$ Ala-Asp-Tyr-Glu-Pro-Ile-Pro-Glu-Glu-Ala-Cha-(D-Glu)-OH, were synthesized by incorporating various nonpolar amino acids at the P3' residue Yaa where Gly was used as a reference amino acid (Table 3). Inhibitors P1054, P1056, P1057, and P1058 with a linear alkyl side chain of L-amino acids (Ala,  $\alpha$ Abu, Nva, and Nle, respectively) at the P3' residue improved the affinity ~3-fold compared to that of the reference inhibitor P836. The inhibitors P1055, P1346, P1075, and P1074 with corresponding D-amino acids (D-Ala, D- $\alpha$ Abu, D-Nva, and D-Nle, respectively), on the other hand, improved the affinity of the inhibitors 25–35-fold, showing ~10-fold preference for D-amino acids over L-amino acids at the P3' residue. Elongation of the linear alkyl side chain did not affect the affinity for both L- and D-amino acids. The inhibitors P1061, P1069, P1059, and P1060 with a branched alkyl side chain of L-amino acids (Val, Tbg, Ile, and Leu, respectively) at the P3' residue improved the affinity 5–20-fold. The branched L-amino acids are visibly preferred over the linear alkyl L-amino acids at the P3' position. The inhibitors P1070, P1073, P1071, and P1072 with a branched side chain of D-amino acids (D-Val, D-Tbg, D-Ile, and D-Leu, respectively) improved the affinity 40–70-fold, significantly higher than the affinity of the corresponding L-amino acids. Introduction of D-Cha with a constrained conformation of the branched alkyl side chains led to less improvement than that of other D-amino acids (P1077; K<sub>i</sub> = (1.5 ± 0.5) × 10<sup>-12</sup> M). A similar lack of improvement was found with sulfur hetero compounds (P1419 with D-Met and P1432 with D-Pen). In addition to an alkyl side chain, aromatic groups were also introduced. The inhibitors P1340, P1339, P1345, P1343, P1344, and P1342 with aromatic L-amino acids (Phg, Phc, Hph, His, Trp, and Thi, respectively) improved the affinity

2–8-fold. The inhibitors P1076, P1078, P1425, P1426, P1427, and P1428 with corresponding D-amino acids (D-Phg, D-Phe, D-Hph, D-His, D-Trp, and D-Thi, respectively) showed better improvement with an 11–250-fold increase in affinity. This indicates a strong preference for aromatic D-amino acids, more specifically for D-Phg (P1076; K<sub>i</sub> = (9.5 ± 0.6) × 10<sup>-14</sup> M) and D-Phe (P1078; K<sub>i</sub> = (1.3 ± 0.5) × 10<sup>-13</sup> M) at the P3' residue. Combining the best substitutions in P1' and P3', Thi and D-Phg, respectively, led to a significantly reduced affinity (P1444; K<sub>i</sub> = (8.6 ± 0.5) × 10<sup>-11</sup> M). There seems to be a conflict in simultaneous insertion of Thi and D-Phg side chains into the S1' and S3' subsites, respectively, even though they are separated by a flexible Gly residue.

**Structures at the Active Site.** Here we report the crystal structures of human  $\alpha$ -thrombin–P628 and –P798 complexes. The P1' residue is Nle in P628 and Thi in P798. The structures of thrombin in these complexes are nearly identical, with an rms deviation of 0.102 for all C<sub>α</sub> atoms. The autolysis loop Glu<sup>146</sup>–Lys<sup>148</sup> is disordered, as reported for other thrombin–bivalent inhibitor complexes that crystallized in the C<sub>2</sub> space group (19, 32, 33). Bbs<sup>H1</sup>, Arg<sup>H2</sup>, and (D-Pip)<sup>H3</sup> occupy the S3, S1, and S2 subsites of thrombin, respectively (Figure 1). The structures of these residues were well defined in the electron density maps. The S3 subsite is an aryl binding pocket surrounded by Tyr<sup>60A</sup>, Trp<sup>60D</sup>, Glu<sup>97A</sup>, Asn<sup>98</sup>, Leu<sup>99</sup>, Ile<sup>174</sup>, Trp<sup>215</sup>, and Glu<sup>217</sup>. The aromatic group of the P3 residue makes favorable stacking interactions with Trp<sup>215</sup>. Replacement of the phenyl group of D-Phe<sup>H1</sup> by a cyclohexyl group of D-Cha<sup>H1</sup> leads to additional contacts with the alkyl side chains of Leu<sup>99</sup> and Ile<sup>174</sup> and provides increased affinity (19). Similarly, the p-tert-butyl group of Bbs<sup>H1</sup> in the current inhibitors interacts with the side chains of Leu<sup>99</sup> and Ile<sup>174</sup>, resulting in 8-fold increased affinity as compared to a corresponding inhibitor with D-Phe<sup>H1</sup> (data not shown). Thus, van der Waals interactions are the key element of the molecular association at the S3 subsite. Sulfonyl oxygen atoms of Bbs<sup>H1</sup> are exposed to the solvent and do not interact with thrombin, suggesting little contribution to the affinity. The Arg<sup>H2</sup> side chain extends into the S1 subsite as reported in thrombin–P498 and –P500 complexes (31). Its guanidinium –NH<sub>3</sub><sup>+</sup> moiety forms a salt bridge (d = 3.0 Å) with the carboxylate group of Asp<sup>189</sup>. The guanidinium =NH of Arg<sup>H2</sup> forms a hydrogen bond (d = 3.0 Å) with Wat<sup>7</sup>, which is further hydrogen-bonded (d = 2.8 Å) to the carbonyl oxygen of (D-Pip)<sup>H3</sup>. (D-Pip)<sup>H3</sup> occupies the S2 subsite of thrombin and adopts a chair conformation. (D-Pip)<sup>H3</sup>, the conformation of which is the same in P498–, P500–, P628–, and P798–thrombin complexes, interacts with Ser<sup>214</sup> O<sup>γ</sup> through Wat<sup>7</sup>. The closest carbonyl carbon of the inhibitor to the catalytic oxygen of Ser<sup>195</sup> is the one of (D-Pip)<sup>H3</sup> with a distance of ~4.5 Å, which assumes the proteolytic stability of these compounds against thrombin.

**Structures at the S1' Subsite.** The specificity of the thrombin S1' site for the small P1' residue was experimentally confirmed by characterizing the reactivity of thrombin with natural and engineered variants of antithrombin III in which Ser<sup>394</sup> at the P1' position is substituted with other amino acids (8, 34–36). The conformation of Lys<sup>60F</sup> was used to explain the substrate specificity of the P1' residue for small amino acids such as Gly, Ala, Val, Ser, Thr, and Asp (Figure 2) (37). In the thrombin–P498 complex, the Lys<sup>60F</sup> is close to the main chain of the inhibitor P498 [the

### P1' and P3' Residues of Thrombin Inhibitors

*Biochemistry*, Vol. 39, No. 9, 2000 2389

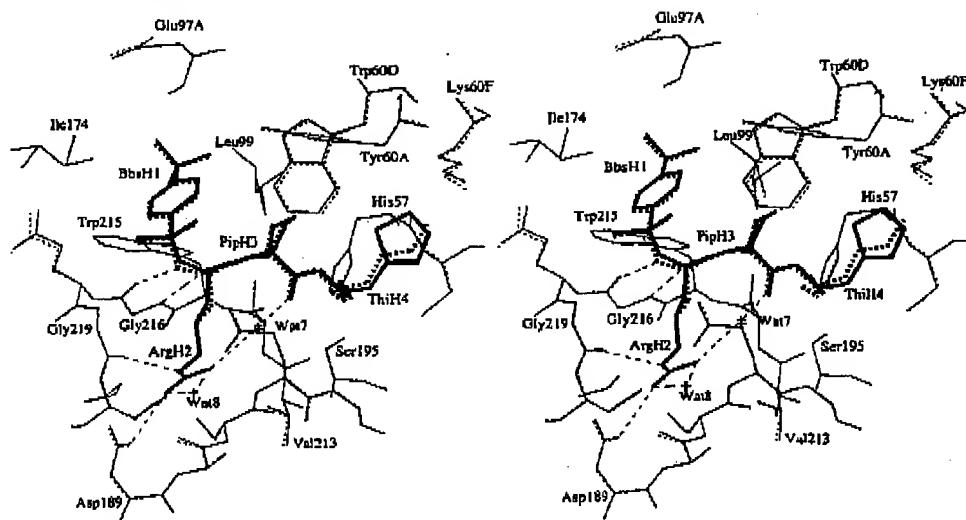


FIGURE 1: Stereoview of the crystal structures of thrombin-P628 and -P798 complexes at the thrombin active site. Thin and thick lines represent the thrombin and P798, respectively, in the thrombin-P798 complex. Thin broken and thick broken lines represent thrombin and P628, respectively, in the thrombin-P628 complex. The dashed lines represent hydrogen bonds.

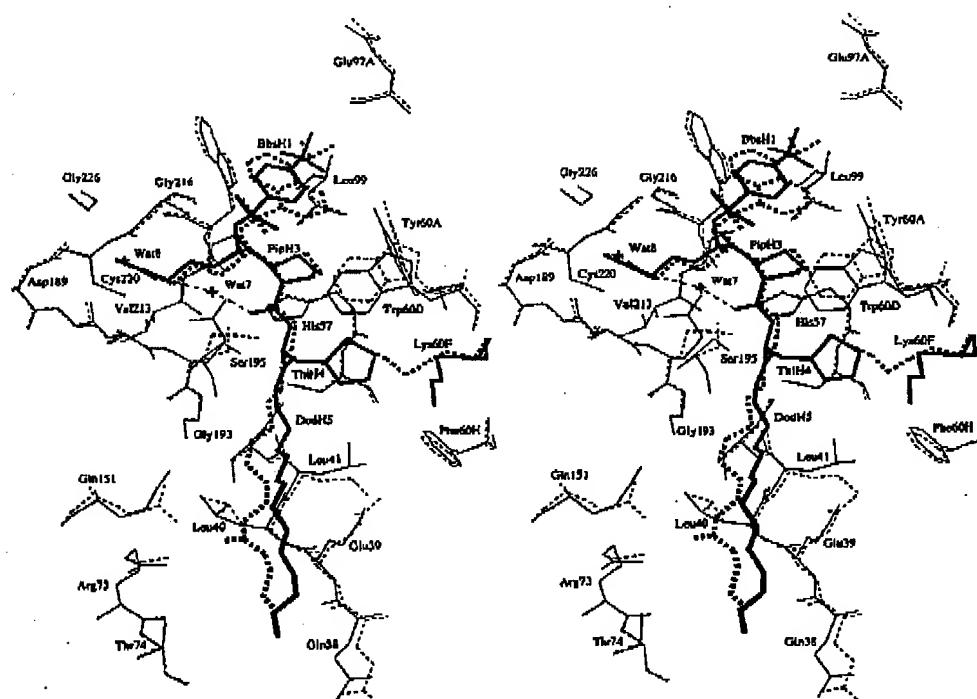


FIGURE 2: Stereoviews of the crystal structures of thrombin-bivalent inhibitor complexes at the S1' subsite. Thin and thick lines represent the thrombin and P798, respectively, in the thrombin-P798 complex. Thin broken and thick broken lines represent thrombin and P500, respectively, in the thrombin-P500 complex for comparison. Dashed lines represent hydrogen bonds.

Ne of Lys<sup>60F</sup> is 4.7 Å away from C<sup>a</sup> of  $\beta$ Ala<sup>H4</sup> and is hydrogen-bonded to the amide nitrogen of  $\beta$ Ala<sup>H4</sup> through Wat<sup>112</sup>. In addition, the side chain orientations of Leu<sup>41</sup>, Glu<sup>39</sup>, and Leu<sup>33</sup> were also different from those of the conformation observed in thrombin-P628 and -P798 complexes. However, this argument does not explain the presence of large amino acids Ile, Leu, and His in some thrombin substrates nor high affinity of the chromogenic and fluorogenic substrates, the significant portion of the affinity of which comes from its bulky nonpolar P1' residue. The electron density of the Lys<sup>60F</sup> side chain in the thrombin-

P498 complex is not very well defined, suggesting flexible conformation and rather poor impediment against a bulky P1' residue. Furthermore, another well-defined conformation of the Lys<sup>60F</sup> side chain was observed by Matthews et al. (38) and St. Charles et al. (39) as well as in thrombin-P628 and -P798 complexes (Figure 2). This conformation of Lys<sup>60F</sup> is well defined in the electron density with an average *B* factor of 28.9 and 33.6 Å<sup>2</sup> for the Lys<sup>60F</sup> side chain atoms of P628 and P798 complexes, respectively. The Lys<sup>60F</sup> Nε is hydrogen-bonded to Wat<sup>114</sup> (*d* = 2.9 Å), which interacts with Glu<sup>39</sup> through Wat<sup>128</sup>. The aliphatic part of the Lys<sup>60F</sup>

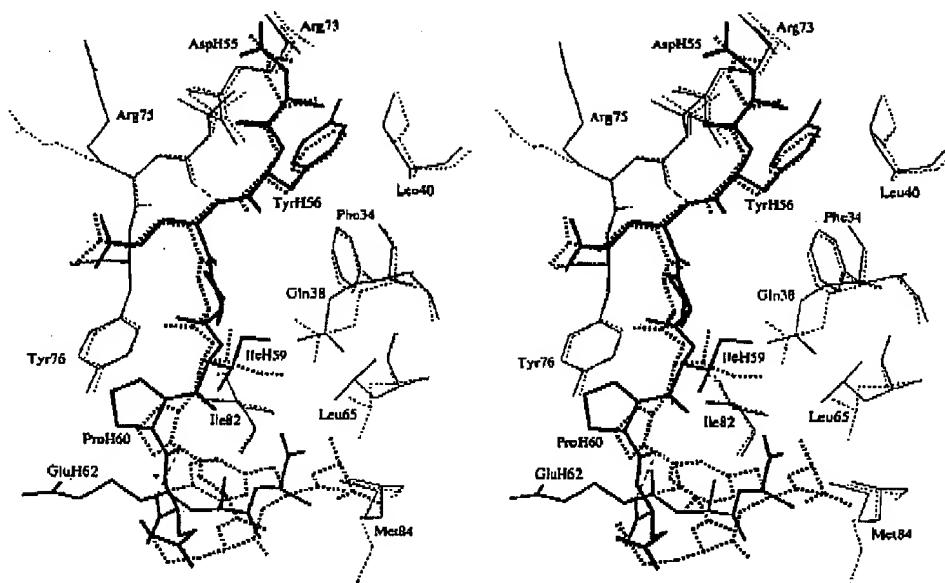


FIGURE 3: Stereoviews of the crystal structures of thrombin-bivalent inhibitor complexes at the FRE. Thin and thick lines represent the thrombin and P798, respectively, in the thrombin-P798 complex. Thin broken and thick broken lines represent thrombin and P500, respectively, in the thrombin-P500 complex for comparison. Dashed lines represent hydrogen bonds.

side chain interacts with the side chains of Leu<sup>41</sup>, Trp<sup>60D</sup>, and Phe<sup>60H</sup> through van der Waals interactions, resulting in a large and deep hydrophobic S1' binding pocket which can accommodate the bulky nonpolar P1' residue. Thus, the S1' subsite is a potential nonpolar binding site that can enhance the affinity of the inhibitors and substrates. The side chain of the P1' residue of P628 or P798 is surrounded by Leu<sup>41</sup>, Cys<sup>42</sup>-Cys<sup>57</sup>, His<sup>57</sup>, Tyr<sup>60A</sup>, Trp<sup>60D</sup>, Phe<sup>60H</sup>, and Lys<sup>60F</sup> as well as D-Pip<sup>13</sup>. The side chain of Thr<sup>14</sup> is extended into the cavity of the S1' subsite. The thiienyl ring stacks against the indole ring of Trp<sup>60D</sup>. The side chain of Nle<sup>14</sup> is also extended into the cavity of the S1' subsite. It forms numerous van der Waals interactions with Leu<sup>41</sup>, Trp<sup>60D</sup>, Phe<sup>60H</sup>, Cys<sup>42</sup>, and Cys<sup>57</sup>. The  $\epsilon$ -methyl group of Nle<sup>14</sup>, especially, interacts with  $\beta$ - and  $\gamma$ -methylene groups of Lys<sup>60F</sup> ( $d = 3.9$  and  $3.8$  Å, respectively).

**Linker Structures.** Crystal structures of thrombin-bivalent inhibitors show that the short linkers occupy the S' subsites (27, 31), whereas long linkers seem to be largely exposed to solvent (27). Slon-Usakiewicz et al. (21) reported that a linker of at least 15 atoms (including the P1' residue) is required for optimum simultaneous bindings to the active site and the FRE. Here we used a 16-atom linker to provide some conformational flexibility. As expected,  $\lambda$ Adod<sup>15</sup> of P628 and P798 lies in the deep cleft along the S' subsites of thrombin (Figure 2). The methylene groups of  $\lambda$ Adod<sup>15</sup> form nonpolar interactions with the side chains of Glu<sup>39</sup> and Glu<sup>192</sup>. The NH group of  $\lambda$ Adod<sup>15</sup> is hydrogen-bonded to Wat<sup>42</sup> ( $d = 3.0$  Å), and the carbonyl group is hydrogen-bonded to Wat<sup>49</sup> ( $d = 2.9$  Å). The carbonyl oxygen of  $\lambda$ Adod<sup>15</sup> forms an intramolecular hydrogen bond with the NH proton of Tyr<sup>156</sup>. The electron density indicates that the N-terminal half of  $\lambda$ Adod<sup>15</sup> is more flexible than the C-terminal half. This moderate conformational flexibility of the linker may optimally locate the P1' and P3' residues into the S1' and S3' subsites.

**FRE Structure.** The structures of the FRE inhibitor segment of P628 and P798 and the thrombin FRE are

essentially identical, although the side chains of Glu<sup>161</sup>, Glu<sup>162</sup>, and Cha<sup>164</sup> as well as the D-Glu<sup>165</sup> residue could not be identified due to the lack of electron density or to the scattered density (Figure 3). The Asp<sup>155</sup> side chain forms a salt bridge ( $d = 2.8$  Å) to Arg<sup>73</sup>. The electron density of Tyr<sup>156</sup> is well defined. Perpendicular aromatic-aromatic stacking was observed between Tyr<sup>156</sup> and Phe<sup>34</sup>. In addition, Met<sup>32</sup>, Leu<sup>40</sup>, Arg<sup>73</sup>, Thr<sup>74</sup>, and Gln<sup>151</sup> side chains form nonpolar interactions with the Tyr<sup>156</sup> side chain. The hydroxy group of Tyr<sup>156</sup> forms a hydrogen bond ( $d = 3.3$  Å) with Arg<sup>73</sup>. The carboxyl group of Glu<sup>157</sup> does not form a salt bridge; however, Glu<sup>157</sup> Oe1 and Oe2 are hydrogen-bonded ( $d = 3.3$  and  $2.7$  Å, respectively) to the backbone NH of Tyr<sup>76</sup> and Wat<sup>49</sup>, respectively. Wat<sup>49</sup> is further hydrogen-bonded ( $d = 2.7$  Å) to the backbone carbonyl oxygen of Pro<sup>158</sup>. The side chains of Glu<sup>161</sup> and Glu<sup>162</sup> are exposed to the solvent with no interaction observed. The electron density of Ile<sup>159</sup> is well defined. The average temperature factor of the entire side chain was 44.5 and 46.3 for P628- and P798-thrombin complexes, respectively. Its side chain is inserted into a hydrophobic pocket surrounded by Phe<sup>34</sup>, Leu<sup>40</sup>, Tyr<sup>76</sup>, Ile<sup>159</sup>, and Cha<sup>164</sup>. The side chain of Ile<sup>159</sup> adopts a conformation with  $\chi_1 = -66^\circ$  pointing  $^4\text{CH}_3$  of Ile<sup>159</sup> toward Arg<sup>67</sup> different from that ( $\chi_1 = 58^\circ$ ) in the hirudin<sup>66-65</sup> FRE inhibitor segment of hirutomin-2 (27). The side chains of Glu<sup>38</sup>, Glu<sup>39</sup>, Leu<sup>41</sup>, Leu<sup>65</sup>, and Met<sup>84</sup> were also in a different orientation (Figure 3). The Pro<sup>160</sup> side chain has van der Waals interactions with Tyr<sup>76</sup>. The residues Glu<sup>161</sup>, Glu<sup>162</sup>, Ala<sup>163</sup> form a  $\text{3}_{10}$  helical turn which is stabilized by a hydrogen bond between the carbonyl oxygen of Pro<sup>160</sup> and the NH proton of Ala<sup>163</sup>, and nonpolar interactions of the Ala<sup>163</sup> side chain with Pro<sup>160</sup> and Ile<sup>159</sup> side chains. This turn brings Cha<sup>164</sup> back to the binding pocket of Ile<sup>159</sup>; however, we did not observe any clear electron density in Cha<sup>164</sup> and (D-Glu)<sup>165</sup> residues, suggesting their conformations are flexible.

**P1' and P3' Residues of Thrombin Inhibitors****CONCLUSIONS**

The linker of bivalent thrombin inhibitors was successfully designed to interact with thrombin S' subsites, resulting in a design of trivalent thrombin inhibitors. Among the trivalent inhibitors with a various nonpolar P1' or P3' residue, incorporation of aromatic D-amino acids at the P3' residue improved the affinity ~200-fold to a  $K_i$  value of  $(9.5 \pm 0.6) \times 10^{-14}$  M with D-Phe and  $(1.3 \pm 0.5) \times 10^{-13}$  M with D-Phe. Incorporation of Nle and Thi at the P1' residue also improved the affinity to a  $K_i$  value of  $(3.8 \pm 0.6) \times 10^{-14}$  or  $(1.7 \pm 0.4) \times 10^{-14}$  M, respectively. Crystal structures of two of the thrombin-inhibitor complexes revealed that incorporation of Nle or Thi accompanies a large conformational change of the Lys<sup>60P</sup> side chain and creates a deep and large nonpolar S1' subsite. This new S1' subsite accepts a large nonpolar P1' residue to improve the affinity of ligands by 2 orders of magnitude. Thus, the design of P1' and P3' residues converted the hirudin-based bivalent inhibitors into trivalent inhibitors, and improved the affinity to the level of natural hirudin ( $K_i = 2.2 \times 10^{-14}$  M) with molecules that are one-third the size of hirudin.

**ACKNOWLEDGMENT**

We thank Drs. Enrico Purisima and Traian Sulea for useful discussion. We also thank Jean Lefebvre for technical assistance.

**REFERENCES**

1. Laudano, A. P., Cottrell, B. A., and Doolittle, R. F. (1983) *Ann. N. Y. Acad. Sci.* 37, 315-329.
2. Vu, T.-K. H., Hung, D. T., Wheaton, V. L., and Coughlin, S. R. (1991) *Cell* 64, 1057-1068.
3. Schechter, I., and Berger, A. (1967) *Biochem. Biophys. Res. Commun.* 27, 157-162.
4. Bode, W., Mayr, I., Baumann, U., Huber, R., Stonc, S. R., and Hofsteege, J. (1989) *EMBO J.* 8, 3467-3475.
5. Blombäck, B., Blombäck, M., Hessel, B., and Iwanaga, S. (1967) *Nature* 215, 1445-1448.
6. Le Bonnec, B. F., Myles, T., Johnson, T., Knight, C. G., Tapparelli, C., and Stone, S. R. (1996) *Biochemistry* 35, 7114-7122.
7. Martin, P. D., Robertson, W., Turk, D., Huber, R., Bode, W., and Edwards, B. F. P. (1992) *J. Biol. Chem.* 267, 7911-7920.
8. Rezaie, A. R., and Olson, S. T. (1997) *Biochemistry* 36, 1026-1033.
9. Maraganore, J. M., Bourdon, P., Jablonsky, J., Ramachandran, K. I., and Fenton, J. W. II (1990) *Biochemistry* 29, 7095-7101.
10. DiMaio, J., Gibbs, B., Munn, D., Lefebvre, J., Ni, F., and Konishi, Y. (1990) *J. Biol. Chem.* 265, 21698-21703.
11. Bourdon, P., Jablonski, J.-A., Chao, B. H., and Maraganore, J. M. (1991) *FEBS Lett.* 294, 163-166.
12. DiMaio, J., Ni, F., Gibbs, B., and Konishi, Y. (1991) *FEBS Lett.* 282, 47-52.
13. DiMaio, J., Gibbs, B., Lefebvre, J., Konishi, Y., Munn, D., Yue, S.-Y., and Horberger, W. (1992) *J. Med. Chem.* 35, 3331-3341.
14. Kline, T., Hammond, C., Bourdon, P., and Maraganore, J. M. (1991) *Biochem. Biophys. Res. Commun.* 177, 1049-1055.
15. Krishnan, R., Tulinsky, A., Vlasuk, G. P., Pearson, D., Valla, P., Bergum, P., Brunck, T. K., and Ripka, W. C. (1996) *Protein Sci.* 5, 422-433.
16. Lombardi, A., Nastri, F., Morte, R. D., Rossi, A., De Rosa, A., Staiano, N., Pedone, C., and Pavone, V. (1996) *J. Med. Chem.* 39, 2008-2017.
17. Skordalakes, E., Elgendi, S., Goodwin, C. A., Green, D., Scully, M. F., Kakkar, V. V., Freyssinet, J.-M., Dodson, G., and Deadman, J. I. (1998) *Biochemistry* 37, 14420-14427.
18. Lombardi, A., De Simone, G., Nastri, F., Galdiero, S., Morte, R. D., Staiano, N., Pedone, C., Bolognesi, M., and Pavone, V. (1999) *Protein Sci.* 8, 91-95.
19. Rehse, P. H., Steinmetzer, T., Li, Y., Konishi, Y., and Cygler, M. (1995) *Biochemistry* 34, 11537-11544.
20. Tsuda, Y., Cygler, M., Gibbs, B. F., Pedyczak, A., Féthière, J., Yue, S.-Y., and Konishi, Y. (1994) *Biochemistry* 33, 14443-14451.
21. Slon-Usakiewicz, J. J., Purisima, E., Tsuda, Y., Sulea, T., Pedyczak, A., Féthière, J., Cygler, M., and Konishi, Y. (1997) *Biochemistry* 36, 13494-13502.
22. Szewczuk, Z., Gibbs, B. F., Yue, S.-Y., Purisima, E., Zdanow, A., Cygler, M., and Konishi, Y. (1993) *Biochemistry* 32, 3396-3404.
23. Szewczuk, Z., Gibbs, B. F., Yue, S.-Y., Purisima, E., and Konishi, Y. (1992) *Biochemistry* 31, 9132-9140.
24. Segel, I. H. (1975) *Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems*, pp 100-160, John Wiley & Sons, New York.
25. Otwinowski, Z. (1993) *Proceedings of the CCP4 Study Weekend: "Data Collection and Processing"*, Jan 29-30, 1993, pp 56-62, SERC Daresbury Laboratory, England (compiled by Sawyer, L., Isaacs, N., and Bailey, S.).
26. Minor, W. (1993) XDISPLAY Program, Purdue University.
27. Zdanov, A., Wu, S., DiMaio, J., Konishi, Y., Li, Y., Wu, X., Edwards, B. F. P., Martin, P. D., and Cygler, M. (1993) *Proteins* 17, 252-265.
28. Jones, T. A., Zhou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A* 47, 110-119.
29. Brunger, A. T. (1993) X-plor version 3.1, Yale University, New Haven, CT.
30. Okuyama, K., and Ohuchi, S. (1996) *Biopolymers* 40, 85-103.
31. Féthière, J., Tsuda, Y., Coulombe, R., Konishi, Y., and Cygler, M. (1996) *Protein Sci.* 5, 1174-1183.
32. Skrzypczak-Jankun, E., Carperos, V. E., Ravichandran, K. G., Tulinsky, A., Westbrook, M., and Maraganore, J. M. (1991) *J. Mol. Biol.* 221, 1379-1393.
33. Qiu, X., Padmanabhan, K. P., Carperos, V. E., Tulinsky, A., Kline, T., Maraganore, J. M., and Fenton, J. W., II (1992) *Biochemistry* 31, 11689-11697.
34. Stephens, A. W., Siddiqui, A., and Hirsh, C. H. W. (1988) *J. Biol. Chem.* 263, 3639-3645.
35. Theunissen, H. J. M., Dijkema, R., Grootenhuis, P. D. J., Swinkels, J. C., DePoorter, T. L., Carati, P., and Visser, A. (1993) *J. Biol. Chem.* 268, 9035-9040.
36. Wendt, S., Janocha, E., Steffens, G. J., and Strassburger, W. (1997) *Protein Eng.* 10, 169-173.
37. Stubbs, M. T., and Bode, W. (1993) *Thrombin Res.* 69, 1-58.
38. Mathews, J. H., Krishnan, R., Costanzo, M. J., Maryanoff, B. E., and Tulinsky, A. (1996) *Biophys. J.* 71, 2830-2839.
39. St Charles, R., Mathews, J. H., Zhang, E., and Tulinsky, A. (1999) *J. Med. Chem.* 42, 1376-1383.

*Biochemistry*, Vol. 39, No. 9, 2000 239115. Krishnan, R., Tulinsky, A., Vlasuk, G. P., Pearson, D., Valla, P., Bergum, P., Brunck, T. K., and Ripka, W. C. (1996) *Protein Sci.* 5, 422-433.16. Lombardi, A., Nastri, F., Morte, R. D., Rossi, A., De Rosa, A., Staiano, N., Pedone, C., and Pavone, V. (1996) *J. Med. Chem.* 39, 2008-2017.17. Skordalakes, E., Elgendi, S., Goodwin, C. A., Green, D., Scully, M. F., Kakkar, V. V., Freyssinet, J.-M., Dodson, G., and Deadman, J. I. (1998) *Biochemistry* 37, 14420-14427.18. Lombardi, A., De Simone, G., Nastri, F., Galdiero, S., Morte, R. D., Staiano, N., Pedone, C., Bolognesi, M., and Pavone, V. (1999) *Protein Sci.* 8, 91-95.19. Rehse, P. H., Steinmetzer, T., Li, Y., Konishi, Y., and Cygler, M. (1995) *Biochemistry* 34, 11537-11544.20. Tsuda, Y., Cygler, M., Gibbs, B. F., Pedyczak, A., Féthière, J., Yue, S.-Y., and Konishi, Y. (1994) *Biochemistry* 33, 14443-14451.21. Slon-Usakiewicz, J. J., Purisima, E., Tsuda, Y., Sulea, T., Pedyczak, A., Féthière, J., Cygler, M., and Konishi, Y. (1997) *Biochemistry* 36, 13494-13502.22. Szewczuk, Z., Gibbs, B. F., Yue, S.-Y., Purisima, E., Zdanow, A., Cygler, M., and Konishi, Y. (1993) *Biochemistry* 32, 3396-3404.23. Szewczuk, Z., Gibbs, B. F., Yue, S.-Y., Purisima, E., and Konishi, Y. (1992) *Biochemistry* 31, 9132-9140.24. Segel, I. H. (1975) *Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems*, pp 100-160, John Wiley & Sons, New York.25. Otwinowski, Z. (1993) *Proceedings of the CCP4 Study Weekend: "Data Collection and Processing"*, Jan 29-30, 1993, pp 56-62, SERC Daresbury Laboratory, England (compiled by Sawyer, L., Isaacs, N., and Bailey, S.).

26. Minor, W. (1993) XDISPLAY Program, Purdue University.

27. Zdanov, A., Wu, S., DiMaio, J., Konishi, Y., Li, Y., Wu, X., Edwards, B. F. P., Martin, P. D., and Cygler, M. (1993) *Proteins* 17, 252-265.28. Jones, T. A., Zhou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A* 47, 110-119.

29. Brunger, A. T. (1993) X-plor version 3.1, Yale University, New Haven, CT.

30. Okuyama, K., and Ohuchi, S. (1996) *Biopolymers* 40, 85-103.31. Féthière, J., Tsuda, Y., Coulombe, R., Konishi, Y., and Cygler, M. (1996) *Protein Sci.* 5, 1174-1183.32. Skrzypczak-Jankun, E., Carperos, V. E., Ravichandran, K. G., Tulinsky, A., Westbrook, M., and Maraganore, J. M. (1991) *J. Mol. Biol.* 221, 1379-1393.33. Qiu, X., Padmanabhan, K. P., Carperos, V. E., Tulinsky, A., Kline, T., Maraganore, J. M., and Fenton, J. W., II (1992) *Biochemistry* 31, 11689-11697.34. Stephens, A. W., Siddiqui, A., and Hirsh, C. H. W. (1988) *J. Biol. Chem.* 263, 3639-3645.35. Theunissen, H. J. M., Dijkema, R., Grootenhuis, P. D. J., Swinkels, J. C., DePoorter, T. L., Carati, P., and Visser, A. (1993) *J. Biol. Chem.* 268, 9035-9040.36. Wendt, S., Janocha, E., Steffens, G. J., and Strassburger, W. (1997) *Protein Eng.* 10, 169-173.37. Stubbs, M. T., and Bode, W. (1993) *Thrombin Res.* 69, 1-58.38. Mathews, J. H., Krishnan, R., Costanzo, M. J., Maryanoff, B. E., and Tulinsky, A. (1996) *Biophys. J.* 71, 2830-2839.39. St Charles, R., Mathews, J. H., Zhang, E., and Tulinsky, A. (1999) *J. Med. Chem.* 42, 1376-1383.

B1992419B

Vol. 27, No. 2, 1967

## BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS

45

Lorkin,

## ON THE SIZE OF THE ACTIVE SITE IN PROTEASES. I. PAPAIN

Israel Schechter and Arieh Berger

The Weizmann Institute of Science, Rehovoth, Israel

1.  
607.

Received March 10, 1967

In order to investigate the size of the active site of papain, we have studied its action on 40 diastereoisomeric peptides of alanine, ranging in length from  $\text{Ala}_2$  to  $\text{Ala}_9$ . The variations observed in the rates of hydrolysis lead to the conclusion that papain, an endopeptidase, has a large active site which extends over about 25 Å and can be divided into 7 "subsites", each accommodating one amino acid residue of the peptide substrate. The subsites are located on both sides of the catalytic site, 4 on the one side and 3 on the other. This suggests that there may be a long "groove" analogous to that found by X-ray analysis in lysozyme (Johnson and Phillips, 1965). The substrate is visualized as fitting into the groove, binding to several subsites of specific geometry (see scheme, Fig. 1).

The active site of an enzyme performs the twofold function of binding the substrate and catalyzing the reaction. The efficiency of these actions determines the

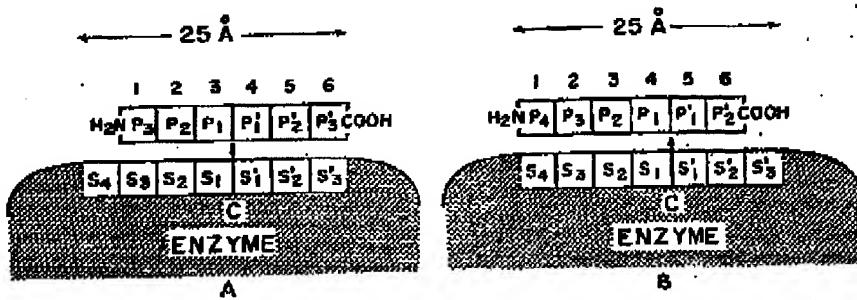


FIG. 1. Schematic representation of two possible enzyme-substrate complexes of papain with a hexapeptide. The active site of the enzyme is composed of 7 "subsites" ( $S_1 - S_4$  and  $S_1' - S_4'$ ) located on both sides of the catalytic site C. The positions,  $P$ , on the hexapeptide substrate are counted from the point of cleavage and thus have the same numbering as the subsites they occupy. Complex A will yield as products two molecules of tripeptide, B one molecule of tetrapeptide and one of dipeptide.

Vol. 27, No. 2, 1967

BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS

Vol. 27, No.

overall activity of the enzyme towards the particular substrate, i.e. determines the specificity of the enzyme. It is therefore possible to obtain information on the active site by studying the kinetics of the enzyme's reactions with different substrates.

An important feature of the active site is its size. It should be possible to "measure" this by using substrates or inhibitors large enough to show up the interactions of the furthermost parts of the binding site. In the present series of investigations on proteolytic enzymes, our approach is to compare the activity of the enzyme towards (a) peptides of increasing length, (b) diastereoisomeric pairs of peptides in which a particular amino acid residue has been replaced by its antipode, and (c) pairs of substrates in which a particular side chain (say a methyl group) has been replaced by another (say an aromatic group). The influence of these changes on reaction rates as a function of distance from the point of cleavage indicates the extent of the active site (Schechter et al., 1965).

Changes in substrate composition may affect either the degree of enzyme-saturation ( $K_m$ ), or the catalytic step ( $k_c$ ), or both. When these parameters cannot be measured separately (e.g.  $K_m$  is too large) the overall efficiency of the enzyme action can be expressed in terms of the proteolytic coefficient C (Irving et al., 1941). This is equal to  $k_c/K_m$ , the rate at infinite substrate dilution.

**MATERIALS AND METHODS.** The synthesis of the alanine peptides used as substrates has already been described (Schechter and Berger, 1966a). Papain, 2x crystallized (Worthington) was used. For reaction conditions see Fig. 3.

The compositions of the reaction mixtures at different times were determined by quantitative paper electrophoresis (Schechter and Berger, 1966b). Two examples are given in Fig. 2. When necessary, the reaction products were isolated and their stereochemical composition determined by further degradation with CP-A and/or IAP (Schechter and Berger, 1966b). For example: the split in  $\text{Ala}_2(\text{DLL}-\text{LL})$  was shown to have occurred at the point indicated, since it gave  $\text{Ala}_2(\text{LL})$  only, and no

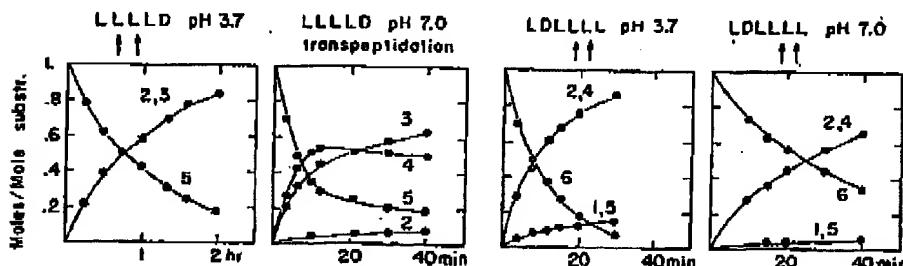


Fig. 2. The course of hydrolysis of alanine peptides by papain. Arrows indicate bonds cleaved; numbers give the size of the peptide.

$\text{Ala}_2(\text{DL})$  showing it was no (L, L) and the split individual.

**RESULTS** are given residue the obs are line always adjace toward the "po and the examp ment o but at

The effect is seen with d be mo striki of the of 10 ment

1. A a D-  $\text{Ala}_2$  assu spec step exis who incr

INS

Vol. 27, No. 2, 1967

## BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS

Ala<sub>2</sub>(DL). Ala<sub>5</sub>(LL-L-LD) gave Ala<sub>3</sub>(LLL 34%; LLD 66%) and Ala<sub>2</sub>(LL 66%; LD 34%). showing splits at both bonds indicated, in a ratio of 2:1. In the case of Ala<sub>5</sub>(LLLLL) it was necessary to synthesize a labelled substrate. Ala<sub>5</sub>(<sup>14</sup>C-L-L-L-L, C<sup>14</sup>) gave Ala<sub>2</sub> (L-L) and Ala<sub>3</sub>(L-L-L) in a ratio of 1:2. From the data obtained, rate constants for the splitting of individual bonds could be assigned unequivocally, the sum of the individual constants in a substrate being equal to the rate of its disappearance.

**RESULTS AND DISCUSSION.** Proteolytic coefficients, C, for cleavages at pH 3.7 are given in Fig. 3 on a log scale<sup>1</sup>. No hydrolysis occurs at bonds involving a D-residue, or at L-L bonds in DL-L sequences (Schechter et al. 1965). We interpret the observed variations in C on the basis of the following assumption: the substrates are lined up on the enzyme in such a way that the CO-NH group being hydrolyzed always occupies the same place (the catalytic site); the amino acid residues occupy adjacent subsites, those towards the NH<sub>2</sub>-end occupying subsites S<sub>1</sub>, S<sub>2</sub> etc., those towards the COOH-end occupying subsites S<sub>1'</sub>, S<sub>2'</sub> etc. For convenience we number the "positions" of the residues in the peptide according to the subsites they occupy, and thus the numbering in a given peptide depends on which bond is split. For example, in the pair Ala<sub>5</sub>(LLLLL) and Ala<sub>5</sub>(DL-L-L-L) the diastereoisomeric replacement of Ala-1 is said to be at P<sub>3</sub> when the bond between Ala-3 and Ala-4 is cleaved - but at P<sub>4</sub> when the bond between Ala-4 and Ala-5 is cleaved (see Fig. 1).

Table I gives the information which can be extracted from Fig. 3 regarding the effect of diastereoisomeric replacements and of elongations by L or D residues. It is seen that the inhibiting effect of D-residues is strongest in S<sub>1</sub> and S<sub>1'</sub>, decreasing with distance from the catalytic point. Subsites on the amino terminal side seem to be more sensitive to replacement (at equal distance) than on the carboxyl side. A striking feature is the cooperative effect of a double replacement on opposite sides of the susceptible bond, as in Ala<sub>5</sub>(DL-L-L) → Ala<sub>5</sub>(L-L-L-L), resulting in a quotient of 100 (J/S, Fig. 3); this is the product of the values for the two single replacements (J/N = 10 and J/Q = 10).

Also in polypeptides, the decrease in hydrolysis rate due to the presence of D-

<sup>1</sup> At pH 7 extensive transpeptidation occurred with peptides which did not contain a D-residue in the N-terminal or next to it. Substrates such as Ala<sub>5</sub>(DL) or Ala<sub>5</sub>(LDL) showed no transpeptidation (see Fig. 2). This can be explained by assuming that the substrate, acting as the acceptor, must occupy the strongly stereospecific subsites S<sub>1'</sub> and S<sub>2'</sub> with its two N-terminal residues during the deacylation step and will therefore fit only if both are of the L-configuration. Evidence for the existence of subsites S<sub>1'</sub>, S<sub>2'</sub> and S<sub>3'</sub> was obtained by Brubacher and Bender (1967) who found that cinnamoyl-papain was deacylated by Gly<sub>n</sub>NH<sub>2</sub> (n = 1 to 3) at rates increasing with n.

Vol. 27, No. 2, 1967

BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS

Vol. 27.

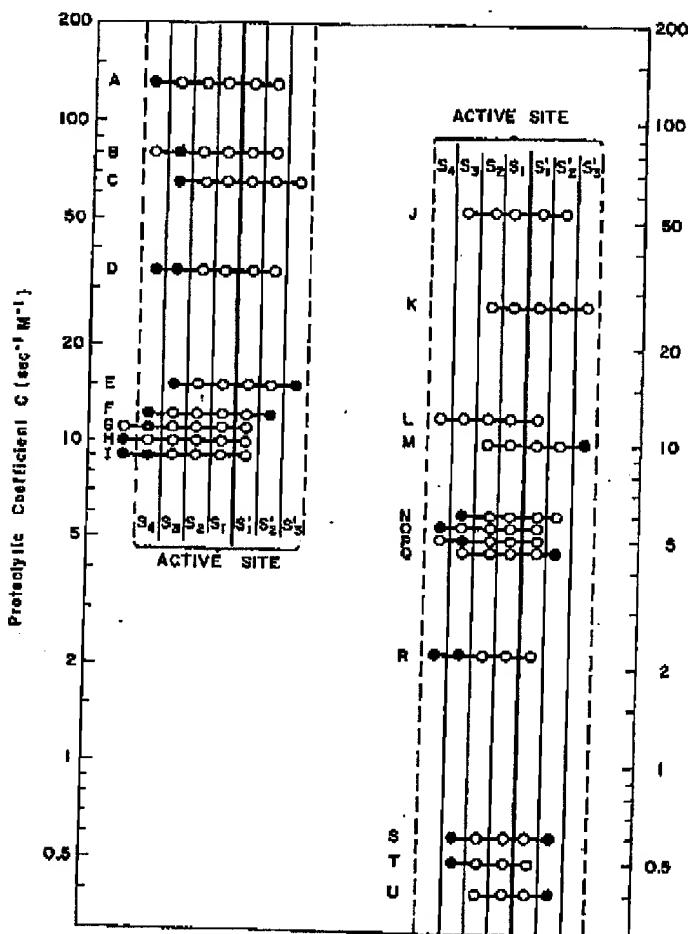


Fig. 3. The proteolytic coefficients C (in  $\text{sec}^{-1} \text{M}^{-1}$ ) for the hydrolysis of individual bonds in a series of alanine peptides at pH 3.7, 36°C. The peptides are drawn schematically at the height corresponding to the C-value for the bond crossing the heavy vertical line (empty circles -○- indicate L-residues, full circles -●- D-residues; the carboxyl is to the right). The subsites ( $S_1$ ,  $S_2$  etc.) occupied by the various residues during the splitting of the bond are indicated. Reaction mixtures: substrate 0.02 M; papain  $1.6 \times 10^{-5}$  M (0.33 mg/ml); mercaptoethanol 0.01 M; EDTA 0.002 M; citric acid 0.025 M;  $\text{Na}_2\text{HPO}_4$  0.016 M.

TABLE

### The new protein values

a)	Dia
	2.4
	2.2
	2.5
b)	Ele
	11
	15
c)	El
	2.
	6.
	2.
	4.

rest  
order  
elge

rep  
due  
con  
S1  
is 1  
lys

TABLE I The Effect of Changes in Substrate Structure on Rates of Hydrolysis

The numbers are the proteolytic quotients (Irving et al., 1941) obtained from the proteolytic coefficients in Fig. 3. The letters in parentheses correspond to the C-values from which the ratio is calculated.

$P_4$	$P_3$	$P_2$	$P_1$	$P_1'$	$P_2'$	$P_3'$
$S_4$	$S_3$	$S_2$	$S_1$	$S_1'$	$S_2'$	$S_3'$
<b>a) Diastereoisomeric replacement D <math>\rightarrow</math> L</b>						
2.4 (B/D)	3.8 (A/D)	$> 10^2$	$> 10^3$	$> 10^3$	12 (A/E)	4.5 (C/E)
2.2 (L/O)	10 (J/N)				10 (J/Q)	2.7 (K/M)
2.5 (P/R)	2.2 (L/P)				9 (N/S)	
	2.6 (O/R)					
	6.3 (Q/S)					
<b>b) Elongation by L-residue</b>						
11 (P/T)	12 (Q/U)	$> 60$	-	-	29 (A/O)	12 (C/N)
16 (B/N)					15 (B/P)	
					15 (D/R)	
					11 (N/T)	
<b>c) Elongation by D-residue</b>						
2.4 (A/J)	1.5 (S/U)	-	-	-	1.2 (S/T)	2.5 (E/N)
6.2 (D/N)	2.4 (C/K)				2.0 (F/O)	
2.2 (F/Q)	1.4 (E/M)					
4.4 (R/T)						

residues follows the pattern of Table I. We found that the DLLD sequences in the ordered copolymer (LDL)<sub>2</sub> are not digested, the DLLLD sequences in (LDLL)<sub>2</sub> are digested slowly (C~1), and the DLLLD sequence in (LLDLL)<sub>2</sub> quite rapidly (C=30).

To sum up, from Table I it can be seen that the effect of diastereoisomeric replacement or of elongation is strongly felt over a range of seven amino acid residues, namely in  $P_1$  to  $P_4$  and  $P_1'$  to  $P_3'$ . This means that the active site of papain corresponds in size to at least seven amino acid residues (subsites  $S_1$  -  $S_4$  and  $S_1'$  -  $S_3'$  in Fig. 1 and Table I). This size - about 25 Å, taking 3.5 Å per residue - is larger than was hitherto suspected. It is roughly as long as the binding site in lysozyme which accommodates 6 sugar units (Phillips, 1966).

If we assume that the active sites in proteases in general are large, it might

Vol. 27, No. 2, 1967

## BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS

Vol. 27,

be possible to explain "unexpected" cleavages observed in proteins (Hill, 1965). It is reasonable to expect that the subsites show, in addition to their stereospecificity, specific interactions with different amino acid side-chains. We have indeed observed this in carboxypeptidase-A (Abramowitz et al., 1967). This would mean that the reactivity of a certain bond in a protein depends not only on the two residues forming this bond (which occupy  $S_1$  and  $S_1'$  during hydrolysis) but also on the nature of the residues in its neighborhood, occupying adjacent subsites. In other words, the enzymes seem to be capable of "recognizing" quite a large portion of a peptide chain.

References

Abramowitz, N., Schechter, I., and Berger, A. (1967), to be published.  
Brubacher, L.J., and Bender, M.L. (1967), to be published.  
Hill, R.L. (1965), *Advan. Protein Chem.* **20**, 37.  
Irving, G.W., Fruton, J.S., and Bergmann, M. (1941), *J. Biol. Chem.* **138**, 231.  
Johnson, L.N., and Phillips, D.C. (1965), *Nature*, **206**, 761.  
Phillips, D.C. (1966), *Scientific American*, **215**, 78.  
Schechter, I., Abramowitz, N., and Berger, A. (1965), *Israel J. Chem.* **3**, 101.  
Schechter, I., and Berger, A. (1966a), *Biochemistry*, **5**, 3362.  
Schechter, I., and Berger, A. (1966b), *Biochemistry*, **5**, 3371.

Received in first temporary und 1966 gly ter fur kin and tia kin vi pr in bl